Alvarado, O., García-Meseguer, R., Javier Ruiz-Pernía, J., Tuñon, I., & Delgado, E. J. (2019). A molecular dynamics study on the role of the protonation state in the biosynthesis of R-PAC by AHAS. *Chemical Physics Letters*, *716*, 247-251. https://doi.org/10.1016/j.cplett.2018.12.039

A molecular dynamics study on the role of the protonation state in the biosynthesis of R-PAC by AHAS

Omar Alvarado[1,2], Rafael García-Meseguer[3,4], José Javier Ruiz-Pernía[4], Iñaki Tuñon[4], Eduardo J. Delgado[*1]

[1] Departamento de Físico-Química, Facultad de Ciencias Químicas, Universidad de Concepción, Concepción, Chile

[2] Departamento de Química, Facultad de Ciencias, Universidad del Bio-Bio, Avenida Collao 1202, Concepción, Chile

[3] School of Mathematics, University of Bristol, Bristol, United Kingdom

[4] Department of Physical Chemistry, Universitat de Valencia, 46100 Burjassot, Spain

Abstract

The effect of the protonation state of the hydroxyl-ethylthiamin diphosphate intermediate, HEThDP, on the enzyme-substrate interactions and their consequences on the biosynthesis of R-phenylacetylcarbinol, R-PAC, by the acetohydroxy acid synthase, AHAS, is addressed by molecular dynamics simulations. It is found that the form of HEThDP, which favors the formation of R-PAC, is that having the 4-aminopyrimidine ring with the N1´ atom protonated and the N4´ atom as aminopyrimidium ion. Under this form both active sites of AHAS have the ability to perform the catalysis, unlike that observed for the other possible protonation states of N1´and N4´ atoms.

[*] E-mail: edelgado@udec.cl

Introduction

R-phenylacetylcarbinol, R-PAC, is an important precursor in the production of a wide variety of drugs with β-adrenergic agonist/antagonist properties. Production of R-PAC using the yeast *Saccharomyces cerevisiae*, glucose and benzaldehyde was the first chiral biotransformation process commercialized [1,2]. In this fermentative route, R-PAC is formed due to the condensation of pyruvate and benzaldehyde catalyzed by the enzyme pyruvate decarboxylase (PDC). The process has low efficiency, substrate toxicity toward living cells and production of important amounts of by-products due to the action of multiple intracellular enzymes. In recent years important efforts have been devoted to the development of processes based on purified catalysts rather than whole cell fermentation [3,4]. It has been found that in addition to catalyze the physiological reactions leading to the formation of (2s)-2-acetohydroxybutyrate and (2S)-2-acetolactate synthase, acetohydroxy acid synthase, AHAS, can catalyze the formation of R-PAC, from the reaction between the intermediate hydroxyl-ethylthiamin diphosphate HEThDP and benzaldehyde (BA), with an enantiomeric specificity over 98 %. Despite knowing the stage of the catalytic cycle in which the non-physiological acceptor substrate benzaldehyde participates, the identification of the protonation state of the 4-aminopyrimidine ring and its consequences on the biosynthesis of R-PAC, has not been clarified so far. This is due to the fact that the R-PAC, formed in steps 3 and 4 of the catalytic cycle of the AHAS competes directly with the formation of the natural product of the enzyme. In a recent article [5], we used a clusterized model to assess the role of the protonation state of the N1´and N4´ atoms of the 4-aminopyrimidine ring of the HEThDP intermediate on the activation barrier. It was observed that the lowest activation barrier is obtained when the residue Glu139 is protonated, i.e. as glutamic acid, and the atom N1´ of 4'-aminopyrimidine ring

deprotonated. The next one in the series corresponds to that having the Glu139 as glutamate and the atom N1´ of 4´-aminopyrimidium ring protonated. Finally, the highest activation barrier corresponded to the tautomer having the Glu139 as glutamate and the atom N1´atom protonated and the N4´atom aminopyrimidium.

The present work addresses the important question of the interplay of the protein environment and the protonation state of the HEThDP intermediate as a determining factor for the enzymatic biosynthesis of R-PAC. Thus, by means of molecular dynamics simulations we studied the effect of the different possible protonation states of the HEThDP-BA complex on the interactions with the active site residues. Our results indicate that the intermediate HEThDP should be with the N1´ atom protonated and the N4´atom as aminopyrimidium, in order to enhance the biosynthesis of R-PAC, in comparison to the physiological products.

Methodology

i) Modeling of the enzyme-substrate complex

The crystal structure of AHAS from *Saccharomyces cerevisiae yeast* in complex with the herbicide chlorimuron ethyl, determined at 2.8 Å (PDB code 1N0H), was used as the starting structure. This structure is composed by two monomers, called A y B from now on. They are composed by 599 and 598 amino acids, respectively. The crystal structure shows two active sites, two molecules of FAD, two atoms of Mg^{2+}, in addition to two molecules of herbicides and 832 crystallization water molecules. The thiazolium and pyrimidine rings that make up the thiamin diphosphate cofactor, ThDP, are joined by a bridge carbon forming the dihedral angles Φ_p = N3-C7'-C5'-C4' and Φ_T = C2-N3-C7'-C5' as shown in Figure 1. When the dihedral angles Φ_p and Φ_T adopt values close to -70° and +95°,

respectively, the cofactor ThDP adopt the active conformation type V, in which the nitrogen atom bound to the pyrimidine ring is oriented close to the carbon C2 of the thiazolium ring.

The cofactor of the monomer A shows the thiazolium ring partially degraded, without the atoms S1 and C2. On the other hand, the cofactor of monomer B does not show loss of atoms in its structure. Besides, between the α and β domains of the two monomers the crystal structure shows the loss of the peptide sequence comprising the residues between Asn271 and Leu276 for the monomer A, and Asn271 and Thr277 for the monomer B [6].

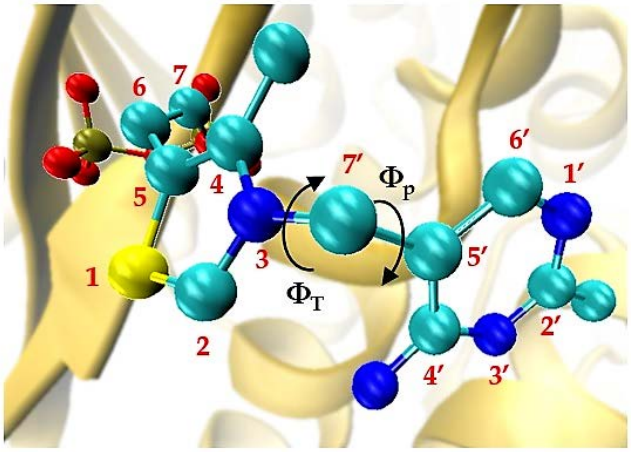


Figure 1. Definition of dihedral angles Φ_p and Φ_T of ThDP.

To rebuild the crystal structure, and to complete the amino acid sequence, the GaussView 5.0.8 [7] and VMD 1.9.2 [8] softwares were used. Briefly, the HEThDP intermediate was positioned by superimposing its ThDP moiety with the ThDP cofactor originally located in such active site. Three possible combinations of the different protonation states of the N1´ and N4´atoms of the HEThDP intermediate were considered, namely, forms AP, APH^+ and IP, as shown in Figure 2.

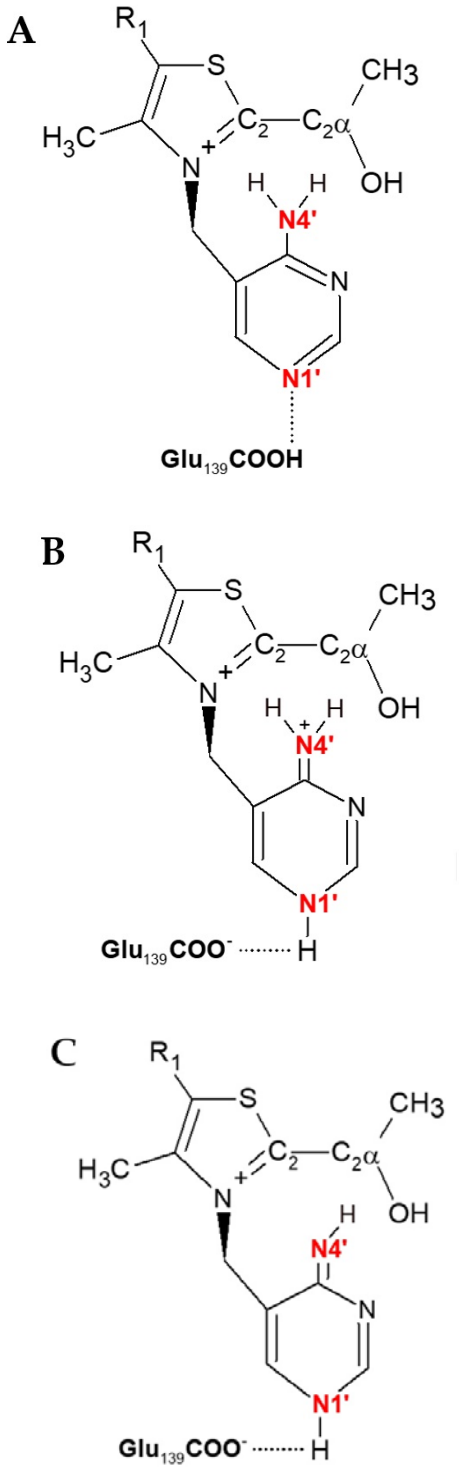


Figure 2. Tautomeric/ionic forms of the HEThDP intermediate: (A) Form AP, (B) Form APH^+, (C) Form IP.

The enzyme-HEThDP-BA complexes were build considering the forms AP, APH^+ and IP of the HEThDP intermediate, generating three complexes labeled, from now on, as CPX-1, CPX-2 and CPX-3, respectively. The protonation states of all ionizable residues were set to the states corresponding to pH 7.0 by using PROPKA 3.0 [9,10]. The enzyme-HEThDP-BA complexes were solvated using a cubic box of water molecules (112 x 102 x 123 $Å^3$) centered in the geometric center of the system. These dimensions of the box ensure that all protein atoms are at least 10 Å away from the edges of the box. The final system for the CPX-1 complex consists of 136384 atoms, of which 18516 are protein atoms, 117564 are water atoms, 6 are sodium ions, 2 are chloride ions, 2 are magnesium ions, 2 are potassium ions incorporated in the crystal structure, and 84 atoms from the FAD cofactor. While, the system for the CPX-2 complex contains 136392 atoms, of which 18514 correspond to protein atoms, 117576 are water atoms, 4 are sodium ions, 2 are magnesium ions, 2 are potassium ions, and 84 atoms from the FAD cofactor. Finally, for CPX-3 the system consists of 136372 atoms, of which 18514 are protein atoms 117552 are water atoms, 8 are sodium ions, 2 are chloride ions, 2 are magnesium ions, 2 are potassium ions, and 84 atoms from the FAD cofactor. For all calculations, protein atoms were described using CHARM36 [11,12], while water molecules were described using TIP3P potentials [13].

ii) Molecular dynamics simulations

Molecular Dynamics (MD) simulations were carried out using the NAMD program [14]. Periodic boundary conditions were imposed and the isothermal-isobaric ensemble (NPT) was used. The temperature (300 K) and pressure (1 atm.) were maintained using NAMD Nosé-Hoover implementations [15,16]. Electrostatic interactions were calculated directly

within a cutoff of 12 Å, whereas long-range electrostatics effects were taken into account by the Particle Mesh Ewald summation method [17]. Van der Waals interactions were treated using a switch function starting at 10 Å and turning off at 12 Å. The parameters and topology files for FAD and the HEThDP-BA complex were obtained using the CGenFF program of CHARMM, [18-20] while the atomic charges were obtained by means of single point calculations of the cofactors employing GAUSSIAN 09 [21] and the CHELPG population analysis.

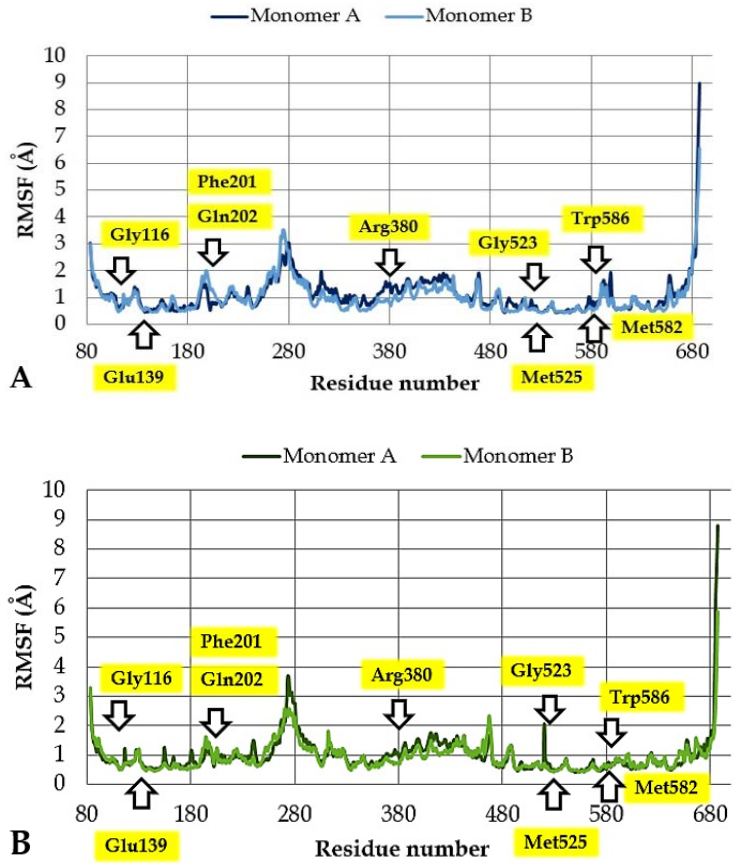
To remove close contacts and the highly repulsive orientation of the initial protein-solvent system, 100000 steps of energy minimization of the water molecules were carried out using the algorithm of conjugated gradients. Afterwards, a new minimization of 100000 steps was carried out for the protein structure of the enzyme to finish with a full minimization of 100000 steps for the total enzyme-solvent system. From the resulting configuration, 100 ps of MD simulations were performed with a time step of 1 fs from 50 K to 400K, with step of 50 K; followed by the system cooling at 300K to equilibrate the system up to the working temperature. Additionally, an equilibrium MD simulation of 1 ns was performed at 300K. This procedure was carried out with the purpose of restoring the positions of the side chains of the residues of the active sites of the systems under study, due to the change made by eliminating the herbicide molecules and the reconstruction of the ThDP cofactor to the complex form. CPX-1, CPX-2 and CPX-3. The temperature changes were made gradually, without breaking the equilibrium of the systems, maintaining the conformations of the secondary and tertiary structures of the systems in addition to important non-binding interactions such as Glu139, Met525 and Gly523. This was done, with the purpose of obtaining initial structures for the DM-MM simulations in which the residues of the active sites were in positions according to the reconstructed complexes. This methodology has

been carried out previously in simulations of molecular dynamics for systems with reconstruction and changes in the molecular structure of the simulated cofactors in order to escape from the local minimum in which the original crystalline structures are found.[22,23] Finally, production simulations of 100 ns with a 2 fs time step were carried out. During these simulations, all intramolecular motions involving hydrogen atoms were frozen by using the SHAKE algorithm [24,25]. Analysis of the trajectories was carried out with VMD.

Results and Discussion

The obtained plot of root-mean-square-deviation (RMSD) for the complete system of both active sites (Figure S1 in Supporting Information) shows that the CPX-2 complex reaches the conformational stability at about 30 ns, while for the other forms a steady increasing value of RMSD is observed until about 70 ns, time at which the values of RMSD for the three complexes are similar. The stability of the secondary and tertiary structure of the protein is assessed by means of the number of hydrogen bonds in the enzyme. The plot of the number of hydrogen bonds in the enzyme for each system along the simulation (Figure S2 in Supporting Information) shows the hydrogens bonds between residues pairs. It is observed a very stable count along the simulation evidencing the stability of the secondary structure for the three complexes, CPX-1, CPX-2 and CPX-3. The root mean square fluctuation (RMSF) results for alpha carbon of key residues located in active sites 1 and 2, after 100 ns simulations, are compiled in the Figure 3. The values show a very stable behavior for the three-protonation forms of the intermediate HEThDP in both active sites 1 and 2, the values vary between 0.4 and 1.2 (Å) and in general remain below 1.0 Å.

It has been suggested in the literature [26,27] that the carbonyl groups of the alternative second substrates make similar contacts in the enzyme active sites as they react with the HEThDP intermediate, therefore the aromatic ring of benzaldehyde and the carboxyl group of a ketoacid must occupy the same general region in the site. Thus, the hydrogen on the new chiral carbon of R-PAC would point in the same direction as the methyl or ethyl group of the acetohydroxy acid [27], since it is postulated that the protonation of the carbonyl oxygen involves the same proton donor in both reactions.



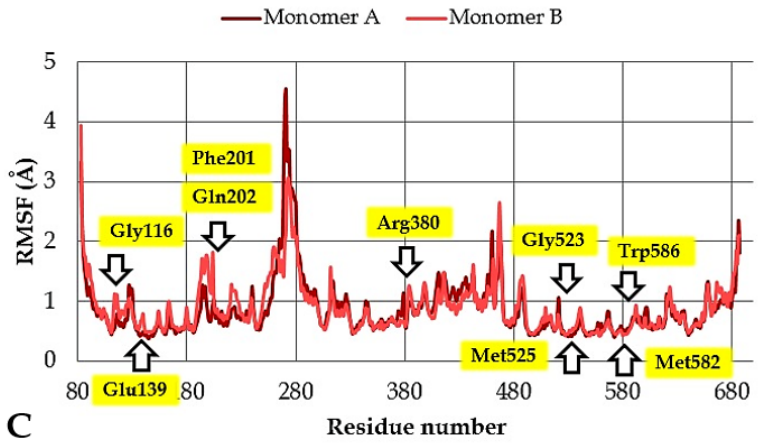
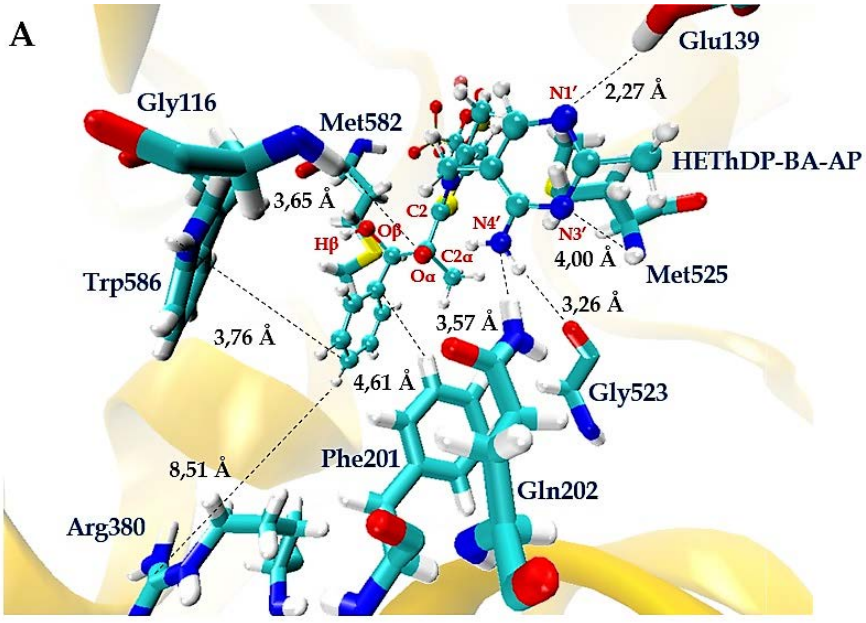


Figure 3. Root mean square fluctuation (RMSF) of key residues for both active sites along the 100 ns of simulation. (A) CPX-1 complex, (B) CPX-2 complex, (C) CPX-3 complex.

Mutagenesis studies [34] show that when Arg380 is replaced by methionine, phenylalanine, glutamine or lysine, the ability to synthesize the physiological product (2S)-2-acetolactate is severely affected, whereas the mutations did not affect the formation of R-PAC, therefore it is postulated that the Arg380 does not play a catalytic role in the formation of R-PAC. Based on these results, it is possible to infer that Arg380 should be located at distances far from the benzaldehyde and HEThDP intermediate. In Figures 4-6, we show the average structures for the 100 ns simulations for both active sites and for the three possible forms of the complexes HEThDP-BA: CPX-1, CPX-2, CPX-3.

It is observed that only the CPX-2 complex meet the conditions mentioned above in both active sites. Thus, the distances from Arg380 to the benzene ring are 10.3 Å and 10.2 Å, for the active site 1 and active site 2, respectively, in agreement with that postulated in literature [27]. While for the other complexes, CPX-1 and CPX-3, this distance is much lower than the above figures, and also show different values for the active sites 1 and 2, in

the same complex. Moreover, it has been proposed [27] that Gly116 would interact by means of hydrogen bonding with the carbonyl oxygen of benzaldehyde, stabilizing the charge density on the oxygen atom and increasing in this way the electrophilicity of the carbonyl carbon. This distance for CPX-2 is 3.17 Å and 2.20 Å for active sites 1 and 2, respectively, compared to the values (4.50 Å, 6.52 Å) and (7.65 Å, 4.73 Å) for the complexes CPX-1 and CPX-3, respectively. On the other hand, Gln202 would interact by means hydrogen bonding with the hydroxyl oxygen of HEThDP, allowing the proton transfer toward the carbonyl oxygen the benzaldehyde [27,28]. The respective values, in the active site 1 and active site 2, are (4.94 Å, 4.71 Å), (2.10 Å, 2.13 Å) and (4.42 Å, 4.92 Å), for the complexes CPX-1, CPX-2 and CPX-3, respectively. All these structural values suggest that this interaction is more favored in the complex CPX-2, compared to complexes CPX-1 and CPX-3. All the above mentioned distances are showed in Figures 4-6.



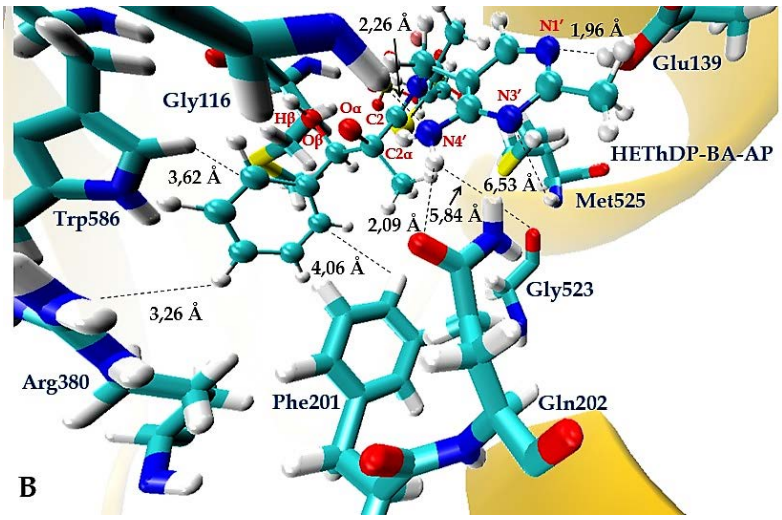
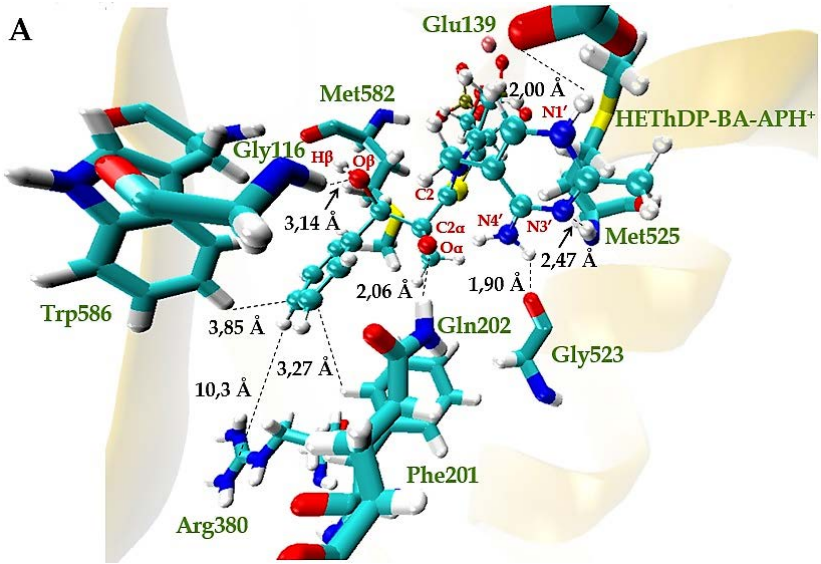


Figure 4. (A) Average structure from the 100 ns simulation for CPX-1 form in the active site 1; (B) Average structure from the 100 ns simulation for CPX-1 form in the active site 2.



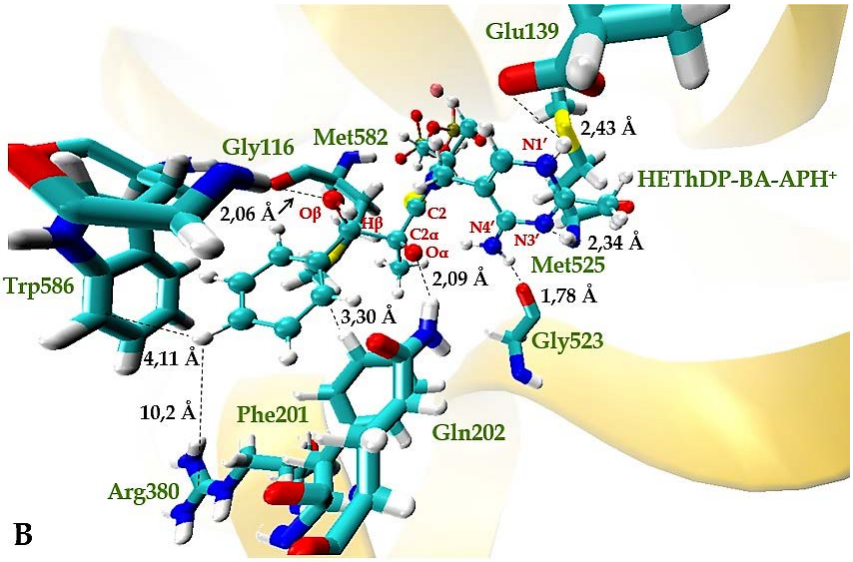
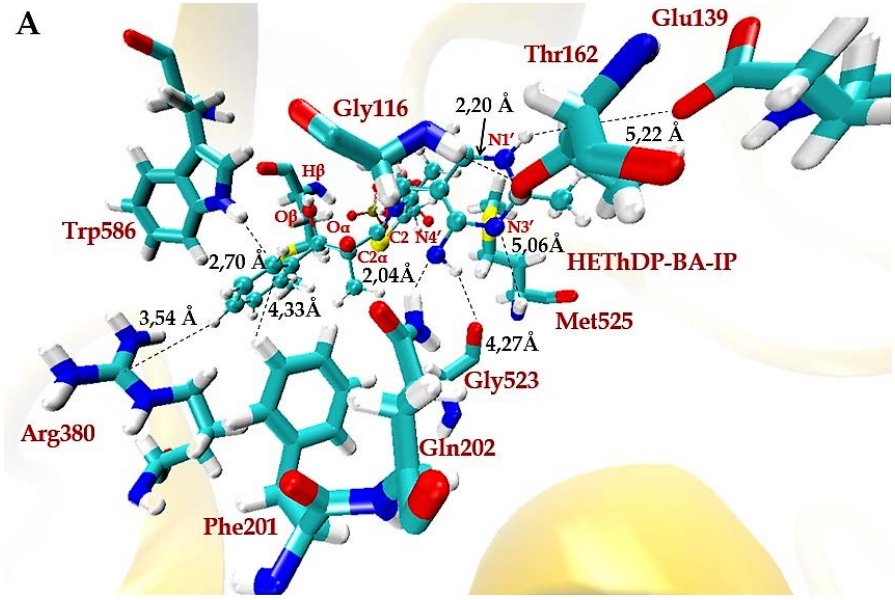


Figure 5. (A) Average structure from the 100 ns simulation for CPX-2 form in the active site 1;(B) Average structure from the 100 ns simulation for CPX-2 form in the active site 2.



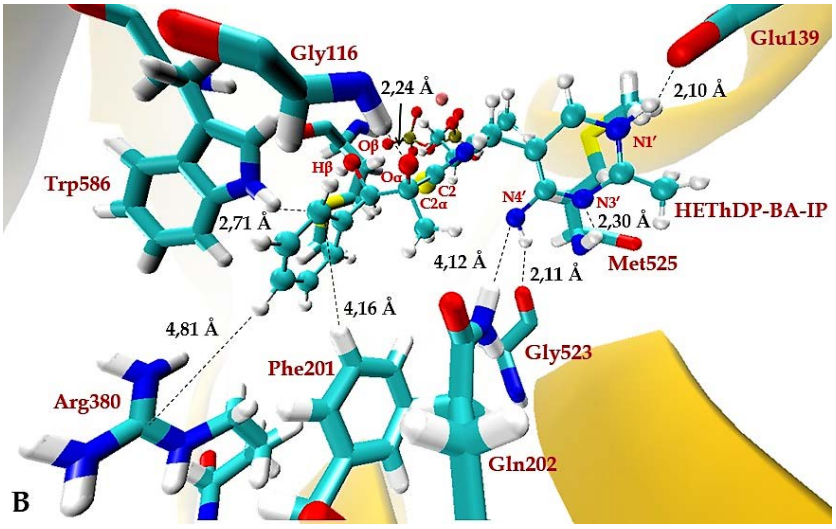


Figure 6. (A) Average structure from the 100 ns simulation for CPX-3 form in the active site 1; (B) Average structure from the 100 ns simulation for CPX-3 form in the active site 2.

On the other hand, it has been suggested that the Met525 play a role purely structural to keep the active conformation V by means of hydrogen bonding interaction with the nitrogen N3´ of the pyrimidine ring of HEThDP [6,29]. The distances values between these two centers, for the active site 1 and active site 2, are (4.06 Å, 6.56 Å), (2.51 Å, 2.37 Å) and (5.00 Å, 2.30 Å) for the complexes CPX-1, CPX-2 and CPX-3, respectively.

Finally, it is necessary to stress that only the complex CPX-2 shows the geometric features required to maintain the interactions with key residues reported in literature. In addition, both active sites show similar characteristics and, in consequence, both active sites are able to carry out the biosynthesis of R-PAC.

Conclusions

The present MD study has been carried out to determine the protonation state of the pyrimidine ring of the intermediate HEThDP during the biosynthesis of R-PAC by AHAS. Our results clearly point that the only protonation state consistent with the structural features reported in the literature is that in which the N1´atom is protonated and the N4´atom is in the form 4'-aminopyrimidium ion, complex CPX-2. In addition, only under this form both active sites show the same right structural features, which provide the ability to perform the catalysis. This condition would grant a maximum efficiency to the enzyme. For the other protonation states of the pyrimidine ring, mixed results are obtained; namely the structural features of both active sites are different each other, although one of them may also show a reasonable agreement with those interactions identified in literature as relevant for the catalysis. Thus, for the complexes CPX-1 and CPX-3, our simulations indicate that only one active site would be able to perform the catalysis reducing the efficiency of the enzyme. On the other hand, having in mind that once the product R-PAC is released, the enzyme must recover its initial condition to initiate the catalytic cycle once more, namely the cofactor ThDP in the form of ylide. The only complex among those considered in this study that would recover the ylide intermediary is the complex CPX-2. In closing, to the best of our knowledge this issue has not been addressed in literature so far, therefore the significance of the study lies on the novelty of the results and they would be of interest for those researchers involved on the development of enzymatic processes based on AHAS for the production of chiral building blocks for bioactive molecules.

Acknowledgement

O.A. acknowledges to Dr. Iñaki Tuñon and at his research group Modeling Chemical Processes in Biological Environments, GIUV2013-011 EFME from physical chemistry department of the University of Valencia, Spain, by the guidance and help delivered during the doctoral internship period in which this research was conducted. Special acknowledge to the Dr. Kirill Zinovjev, Dr. Juan Luis Pascual-Ahuir and the Dr. Manuel Delgado and Dr. Sergio Marti from department of physical and analytical chemistry of University Jaume I, Spain. O.A. acknowledges the financial support from CONICYT for the Doctoral scholarship CONICYT-PCHA/Doctorado Nacional/2013-21130346. from the Chilean Government; the network of doctoral programs REDOC.CTA from Universidad de Concepción, Chile, is also acknowledged. E.J.D acknowledges FONDECYT grant 1170091. I.T., J.J.R.P. and R.G.M. acknowledge financial support from the Ministerio de Economía y Competitividad (Spain) and FEDER funds (project CTQ2015-66223-C2).

References

[1] C. Andreu, M.L. del Olmo, Applied Microbiology and Biotechnology 98 (2014) 5901.

[2] A.L. Oliver, B.N. Anderson, F.A. Roddick, Advances in Microbial Physiology 41 (1999) 1.

[3] H.S. Shin, P.L. Rogers, Applied Microbiology and Biotechnology 44 (1995) 7.

[4] P.L. Rogers, H.S. Shin, B. Wang, Advances in Biochemical Engineering Biotechnology 56 (1997) 33.

[5] O. Alvarado, I. Lizana, G. Jana, I. Tunon, E. Delgado, Chemical Physics Letters 677 (2017) 30.

[6] S.S. Pang, L.W. Guddat, R.G. Duggleby, Journal of Biological Chemistry 278 (2003) 7639.

[7] D.D. Roy, A.K. Todd, M.M. John, Wallingford, CT, USA., Gaussian, Inc., Wallingford, CT, USA., 2009.

[8] W. Humphrey, A. Dalke, K. Schulten, Journal of Molecular Graphics & Modelling 14 (1996) 33.

[9] M.H.M. Olsson, C.R. Sondergaard, M. Rostkowski, J.H. Jensen, Journal of Chemical Theory and Computation 7 (2011) 525.

[10] C.R. Sondergaard, M.H.M. Olsson, M. Rostkowski, J.H. Jensen, Journal of Chemical Theory and Computation 7 (2011) 2284.

[11] R.B. Best, X. Zhu, J. Shim, P.E.M. Lopes, J. Mittal, M. Feig, A.D. MacKerell, Journal of Chemical Theory and Computation 8 (2012) 3257.

[12] B.R. Brooks, C.L. Brooks, A.D. Mackerell, L. Nilsson, R.J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A.R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R.W. Pastor, C.B. Post, J.Z. Pu, M. Schaefer, B. Tidor, R.M. Venable, H.L. Woodcock, X. Wu, W. Yang, D.M. York, M. Karplus, Journal of Computational Chemistry 30 (2009) 1545.

[13] W.L. Jorgensen, J. Chandrasekhar, J.D. Nadura, The Journal of Chemical Physics 79 (1983) 926.

[14] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kale, K. Schulten, Journal of Computational Chemistry 26 (2005) 1781.

[15] G.J. Martyna, D.J. Tobias, M.L. Klein, Journal of Chemical Physics 101 (1994) 4177.

[16] S.E. Feller, Y.H. Zhang, R.W. Pastor, B.R. Brooks, Journal of Chemical Physics 103 (1995) 4613.

[17] T. Darden, D. York, L. Pedersen, Journal of Chemical Physics 98 (1993) 10089.

[18] K. Vanommeslaeghe, A.D. MacKerell, Journal of Chemical Information and Modeling 52 (2012) 3144.

[19] K. Vanommeslaeghe, E.P. Raman, A.D. MacKerell, Journal of Chemical Information and Modeling 52 (2012) 3155.

[20] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, A.D. MacKerell, Journal of Computational Chemistry 31 (2010) 671.

[21] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A. Montgomery Jr., J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J.M. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, Ö. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, D.J. Fox, Gaussian, Inc., Wallingford, CT, USA., 2009.

[22] I. Lizana, G.A. Jana, E.J. Delgado, Journal of Chemical Information and Modeling 55 (2015) 1640.

[23] L. Sanchez, G.A. Jana, E.J. Delgado, Journal of Computational Chemistry 35 (2014) 488.

[24] J. Ryckaert, G. Ciccotti, H.J.C. Berendsen, Journal of Computational Physics 23 (1977) 327.

[25] S.A. Adcock, J.A. McCammon, Chemical Reviews 106 (2006) 1589.

[26] S. Engel, M. Vyazmensky, S. Geresh, Z. Barak, D.M. Chipman, Biotechnology and Bioengineering 83 (2003) 833.

[27] S. Engel, M. Vyazmensky, M. Vinogradov, D. Berkovich, A. Bar-Ilan, U. Qimron, Y. Rosiansky, Z. Barak, D.M. Chipman, Journal of Biological Chemistry 279 (2004) 24803.

[28] M. Vyazmensky, A. Steinmetz, D. Meyer, R. Golbik, Z. Barak, K. Tittmann, D.M. Chipman, Biochemistry 50 (2011) 3250.

[29] J.A. McCourt, R.G. Duggleby, Amino Acids 31 (2006) 173.